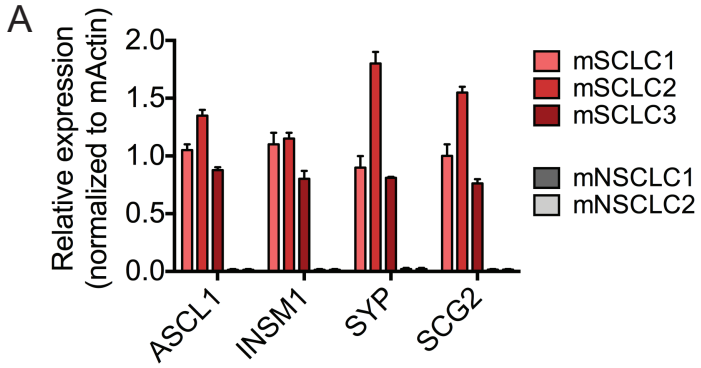
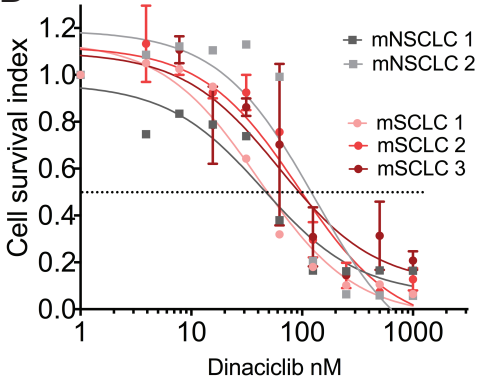
A

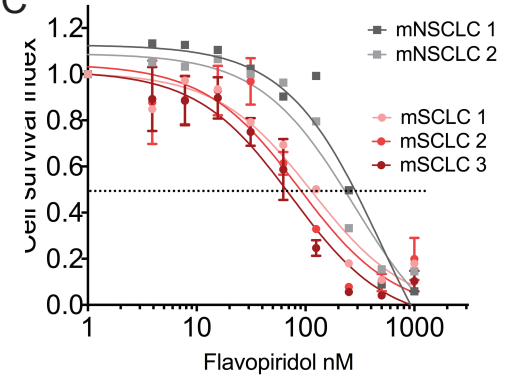


Target	Mean CT (SD) mSCLC1	mNSCLC1
ASCL1	24.47 (0.079)	34.56 (0.671)
INSM1	25.45 (0.070)	34.86 (0.142)
ACTIN	24.98 (0.075)	24.58 (0.069)

B

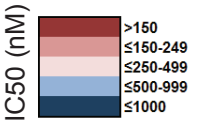


C

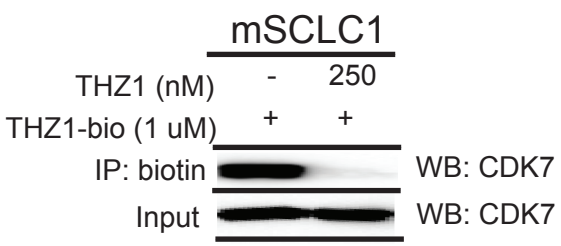


D

TARGET(S)	INHIBITOR	mSCLC mNSCLC
CDK1, CDK2, CDK9	AZD5438	
CDK2, CDK5, CDK7	ROSCOVITONE	
CDK7 (non-covalent)	BS-181	
CDK1, CDk2, CDK9	PHA-767491	
CDK2, CDK6, CDK7, CDK9	BMS-387032	
CDK2, CDK5, CDK7	PHA-793887	
CDK1, CDK2, CDK5, CDK7	FLAVOPIRIDOL	
CDK1,CDK2, CDK5, CDK9	DINACICLIB	
CDK7 (covalent)	THZ1	
CDK7 (non-covalent)	THZ1-R	



E



F

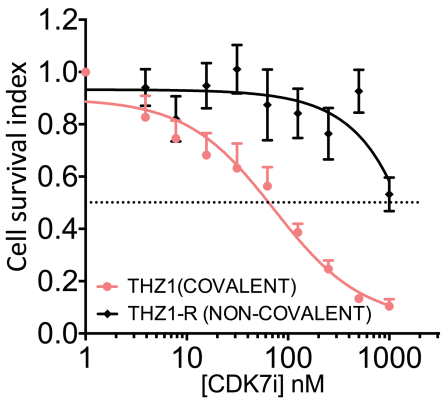


Figure S1. Related to Figure 1. Sensitivity of mSCLC and mNSCLC cells to THZ1, THZ1-R and a panel of commercially available pan-CDK inhibitors.

A) Quantitative PCR (qPCR) to detect expression of classical neuroendocrine (SCLC-specific) markers in murine SCLC (mSCLC1, mSCLC2, mSCLC3: $Rb^{L/L}$;$p53^{L/L}$) and murine NSCLC (mNSCLC) cell lines (mNSCLC1: $Kras^{+/LSL\text{-}G12D}$;$p53^{L/L}$, mNSCLC2: $Kras^{+/LSL\text{-}G12D}$;$p53^{L/L}$;$Lkb1^{L/L}$). Target gene expression is normalized to *Actin* expression and 'C7' serve as reference sample set to 'Relative expression' index 1). Data are presented as mean +/- SEM.

B) Murine SCLC cell lines and murine NSCLC cell lines were incubated with increasing doses of Dinaciclib and IC50s were established from measuring cell viability after two cellular doubling times (5 days for mSCLC and 3 days for mNSCLC cells). Data are presented as mean +/- SEM.

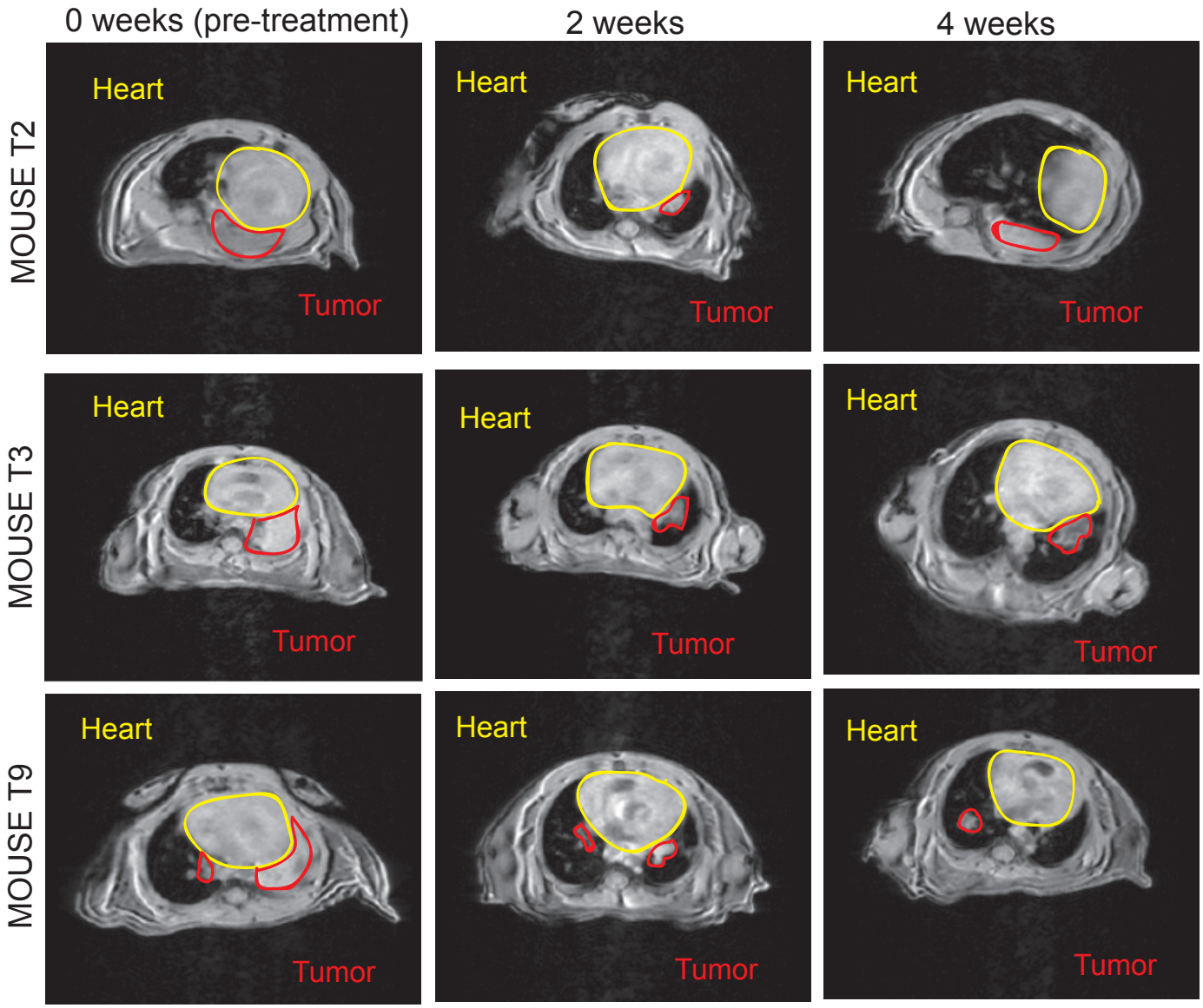
C) Same as A) but with Flavopiridol. Data are presented as mean +/- SEM.

D) 'IC50 heatmap' summary of a panel of pan-CDK inhibitors (including Dinaciclib and Flavopiridol as plotted in Figure S1B-C), THZ1 (also Figure 1E, S1F), and THZ1-R (also Figure S1F) in mSCLC and mNSCLC cell lines. Heatmap key is shown on right.

E) Analysis of target (CDK7) engagement in mSCLC1 cells following THZ1 (250 nM) treatment or control (-) treatment. Cell lysates were chased with THZ1-biotin (THZ1-bio) and pulled down with anti-biotin antibody and analyzed for CDK7 expression by Western blotting. 'Input' represents cell lysate with no added THZ1-bio, which was a control analyzed for CDK7 expression by Western blotting.

F) Murine SCLC cell lines (mSCLC1, mSCLC2, mSCLC3) were incubated with increasing doses of THZ1 or THZ1-R and IC50s were established from measuring cell viability after two cellular doubling times (5 days for mSCLC). Data are presented as mean +/- SEM.

A


0 weeks (pre-treatment)
2 weeks
4 weeks
MOUSE T2
MOUSE T3
MOUSE T9
Heart
Tumor

SUPPLEMENTAL DATA

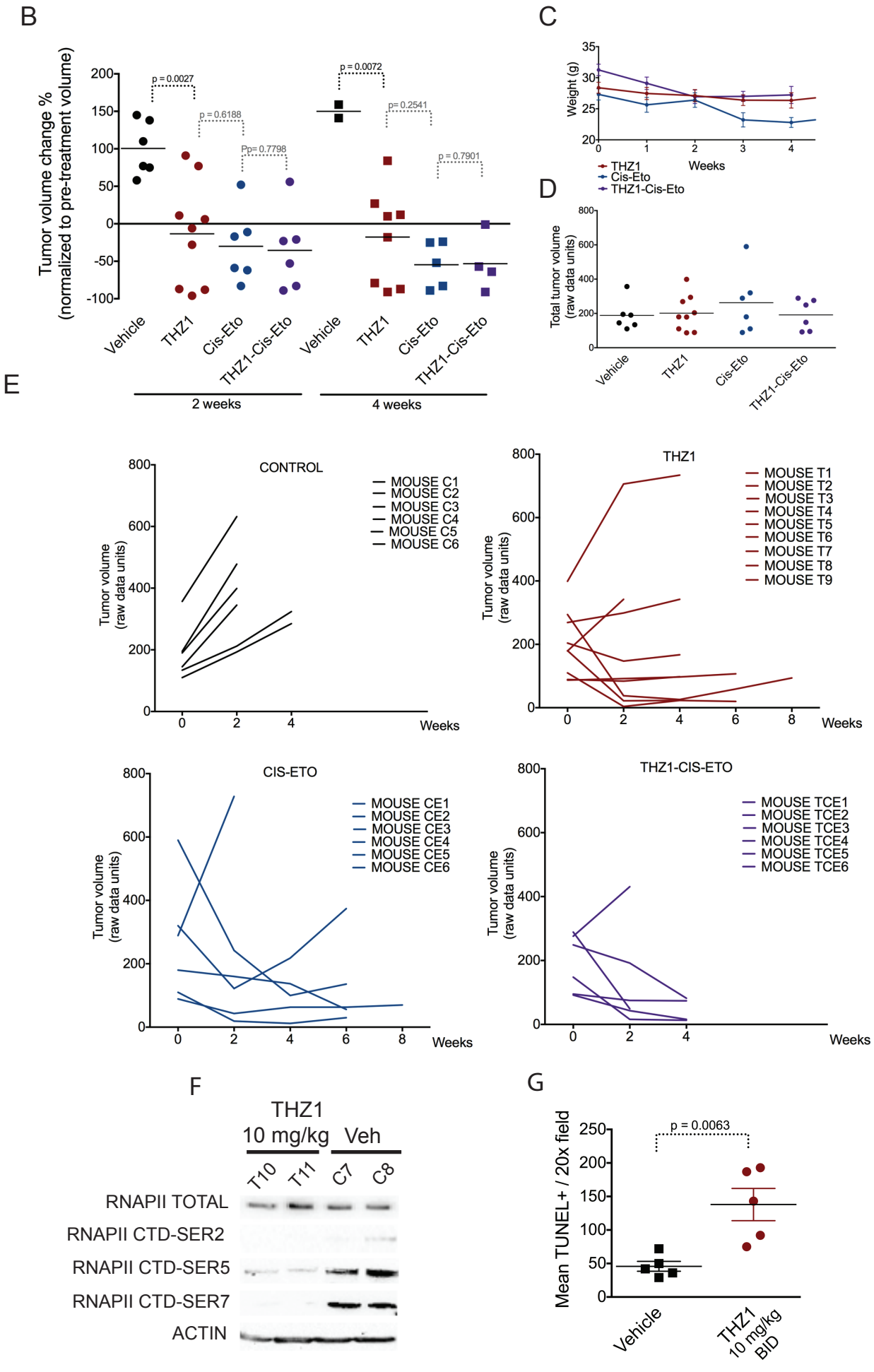


Figure S2. Related to Figure 2. Treatment response in THZ1-treated RP mice

A) MRI of thorax region of THZ1-treated RP mice ($Rb^{L/L}$;$p53^{L/L}$) at 0 (pre-treatment), 2 and 4 weeks experiencing partial response (n=3) with outline of heart (yellow) and tumor (red).

B) Tumor volume changes (%) as measured by MRI in vehicle-treated (control), THZ1-treated, Cisplatin-Etoposide-treated and THZ1-Cisplatin-Etoposide-treated mice at 2 and 4 weeks as normalized to pre-treatment (week 0) tumor volume. p-values between different treatment cohorts were calculated using Students t-test.

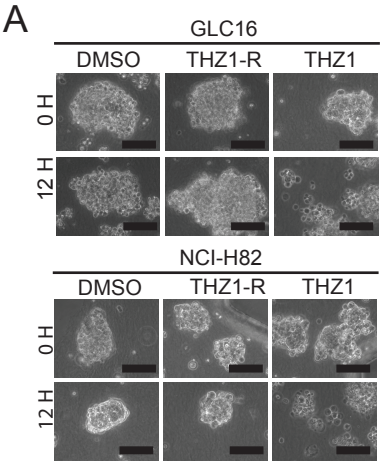
C) Weights of mice on treatment (see panel A))

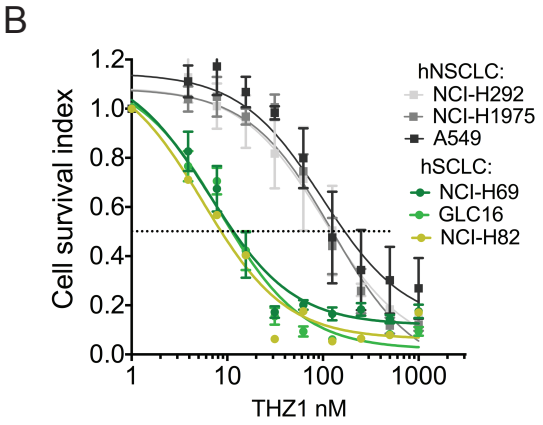
D) Raw tumor volume data (as obtained from 3D Slicer software) of pre-treatment MRIs (week 0) of mice on treatment (see panel A))

E) Raw tumor volume data (as obtained from 3D Slicer software) of vehicle-treated (control), THZ1-treated, Cisplatin-Etoposide-treated and THZ1-Cisplatin-Etoposide-treated mice (raw units used for calculation of tumor volume change % in panel B)).

F) Western blotting to detect RNAPII-CTD phosphorylation in protein lysate from tumor tissue isolated from mice treated with either THZ1 or vehicle for 72 hours. ACTIN serves as loading control.

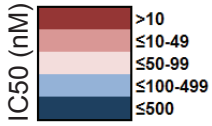
G) TUNEL positive (+) cells in SCLC lung tumors (n=3) and liver metastases (n=1) isolated from RP mice after treatment with THZ1 or vehicle for 72 hours. Quantification was done by manual counting of TUNEL positive nuclei in TUNEL stained tissue sections in 20x magnification field. Mean value of TUNEL+ cells was established from observing 3 fields per tissue section.

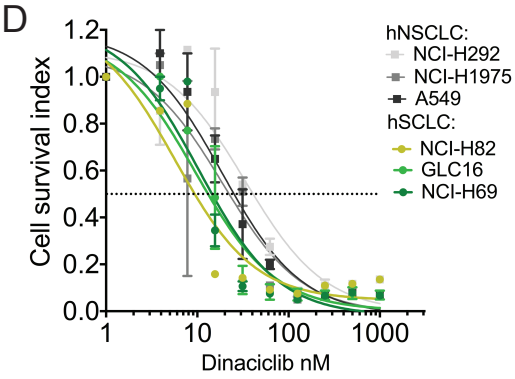


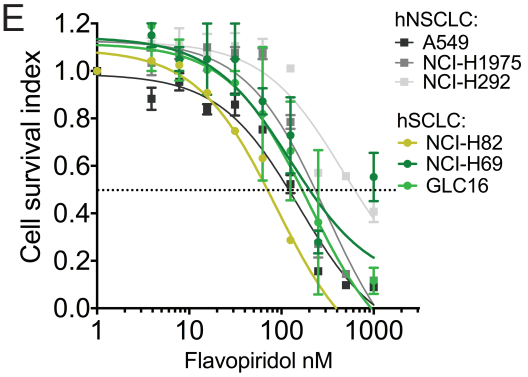


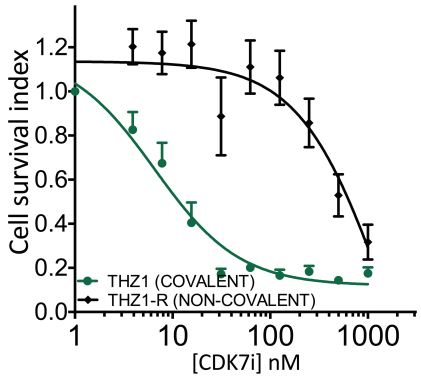
C

TARGET(S)	INHIBITOR	hSCLC	hNSCLC
CDK1, CDK2, CDK9	AZD5438		
CDK2, CDK6, CDK7, CDK9	BMS-387032		
CDK2, CDK5, CDK7	PHA-793887		
CDK7 (non-covalent)	BS-181		
CDK2, CDK5, CDK7	ROSCOVITONE		
CDK1, CDk2, CDK9	PHA-767491		
CDK1, CDK2, CDK5, CDK7	FLAVOPIRIDOL		
CDK1,CDK2, CDK5, CDK9	DINACICLIB		
CDK7 (covalent)	THZ1		
CDK7 (non-covalent)	THZ1-R		









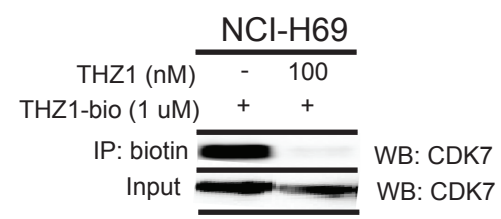


Figure S3. Related to Figure 3. Sensitivity of hSCLC and hNSCLC cells to THZ1, THZ1-R and pan-CDK inhibitors

A) SCLC cells (GLC16 and NCI-H82) grown in 3D matrix culture were treated with DMSO (control), 100 nM THZ1-R or 100 nM THZ1 for 12 hours. Representative phase-contrast microscope images (20x field) from 0 and 12 hours after treatment start with DMSO, THZ1-R or THZ1. Black scale bar represent 50μm.

B) hSCLC cell lines NCI-H69, GLC16, and NCI-H82 and hNSCLC cell lines A549, H1975, and H292 were incubated with increasing doses of THZ1 and IC50s were established from measuring cell viability after two cellular doubling times (5 days for hSCLC and 3 days for hNSCLC cells). Data are presented as mean +/- SEM.

C) Summary represented as an 'IC50 heatmap' of a panel of pan-CDK inhibitors (including Dinaciclib and Flavopiridol as plotted in Figure S4D-E, THZ1 (also Figure 3AB, S4A-B), and THZ1-R (also Figure S4B) in hSCLC (NCI-H69, GLC16, NCI-H82) and hNSCLC (H1975, H292, A549) cell lines. Heatmap key is shown on right.

D) hSCLC cell lines NCI-H69, GLC16, and NCI-H82 and hNSCLC cell lines A549, H1975, and H292 were incubated with increasing doses of Dinaciclib and IC50s were established from measuring cell viability after two cellular doubling times (5 days for hSCLC and 3 days for hNSCLC cells). Data are presented as mean +/- SEM.

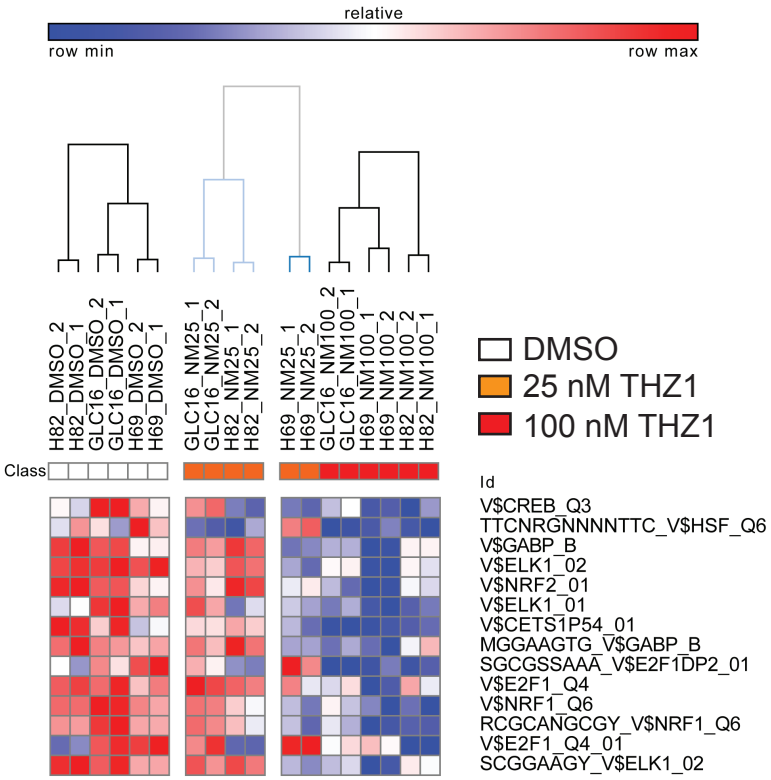
E) Same as D) but with Flavopiridol. Data are presented as mean +/- SEM.

F) hSCLC cell lines NCI-H69, GLC16, and NCI-H82 were incubated with increasing doses of THZ1 or THZ1-R and IC50s were established from measuring cell viability. Data are presented as mean +/- SEM.

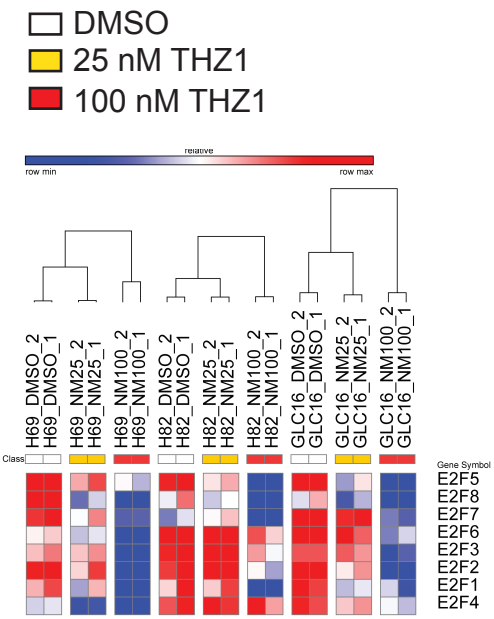
G) Analysis of target (CDK7) engagement in NCI-H69 cells following THZ1 (100 nM) treatment or control (-) treatment. Cell lysates were chased with THZ1-biotin (THZ1-bio) and pulled down with anti-biotin antibody and analyzed for CDK7 expression by Western

blotting. ‘Input’ represents cell lysate with no added THZ1-bio, which was a control analyzed for CDK7 expression by Western blotting.

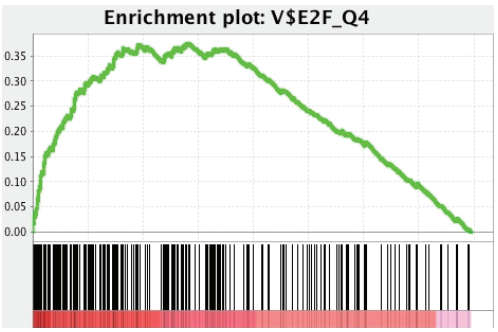
A



B

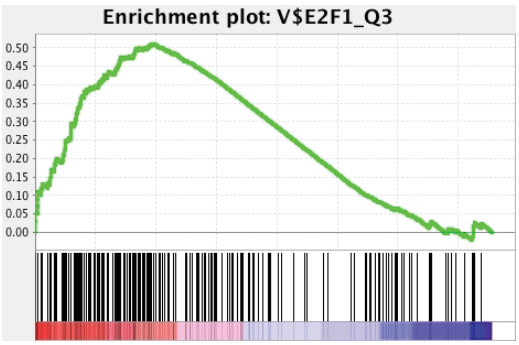


C



THZ1 versus DMSO

D



SCLC versus NSCLC

E

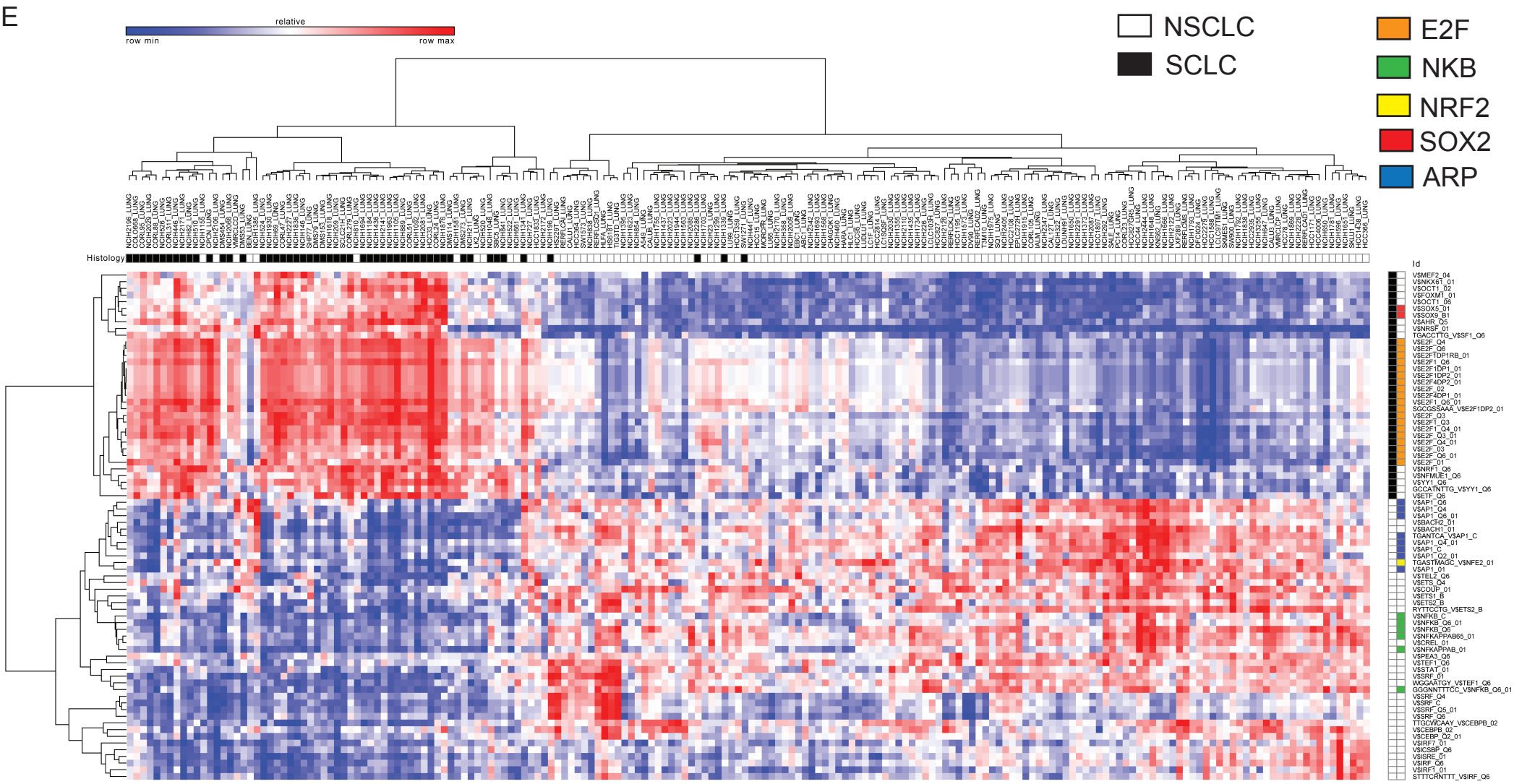

relative
row min
row max
NSCLC
SCLC
E2F
NKB
NRF2
SOX2
ARP
Histology
Id

Figure S4. Related to Figure 4. EF2 gene family in SCLC biology

A) Transcription Factor Binding Sites (TFBS) gene sets identified by GSEA as significantly enriched in 100 nM THZ1-treated cells versus DMSO-treated cells.

B) Heatmap showing expression of *E2F* family genes in NCI-H69, GLC16, and NCI-H82 cells treated with DMSO (control) or 25 nM or 100 nM THZ1 for 6 hours.

C) Enrichment plots of E2F-regulated transcripts THZ1-downregulated transcripts (100 nM) versus DMSO (FDR<0.05).

D) Enrichment plots of Transcription Factor Binding Sites (TFBS) in SCLC versus NSCLC cell lines using Cancer Cell Line Encyplodia data (FDR<0.05)

E) TFBS gene sets identified by GSEA significantly enriched in SCLC (black) versus NSCLC (white). E2F: orange; SOX: red; NFKB: green; NRF2: yellow; AP1: blue.

Table S1. Related to Figure 4. **Summary of Log2 fold gene expression changes in THZ1-treated SCLC cells**

Table S1: Summary of Log_2 fold changes in gene expression in NCI-H69, GLC16, and NCI-H82 cells exposed to 25 nM or 100 nM THZ1 (6 hours) as normalized to control-treated (DMSO) cells.

Cell line	NCI-H69		GLC16		NCI-H82	
[THZ1] nM	25	100	25	100	25	100
% Total downregulation	70	94	82	96	10	82
% 2-fold downregulation	6	27	0.3	25	0.2	13

Table S2. Related to Figure 4. Average log2 expression values per transcript in gene expression analysis if THZ1-treated SCLC cells

Provided as an excel file

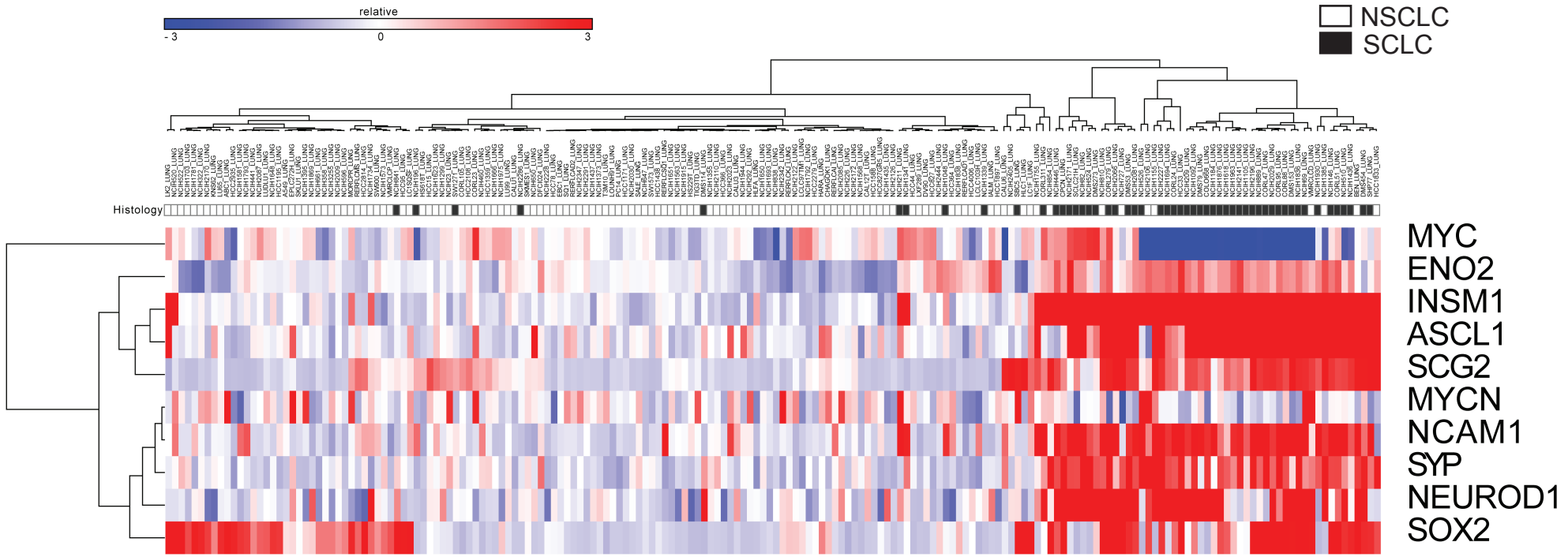

relative
- 3
0
3
NSCLC
SCLC
Histology
MYC
ENO2
INSM1
ASCL1
SCG2
MYCN
NCAM1
SYP
NEUROD1
SOX2

Figure S5. Related to Figure 5. MYC family and lineage-specific transcription factors in SCLC biology.

Heatmap showing expression levels of MYC-family, neuroendocrine and embryonic genes in SCLC and NSCLC cells from the Cancer Cell Line Encyplodia dataset.

Table S3. Related to Figure 5. Enhancer analysis of SCLC cells

Provided as an excel file

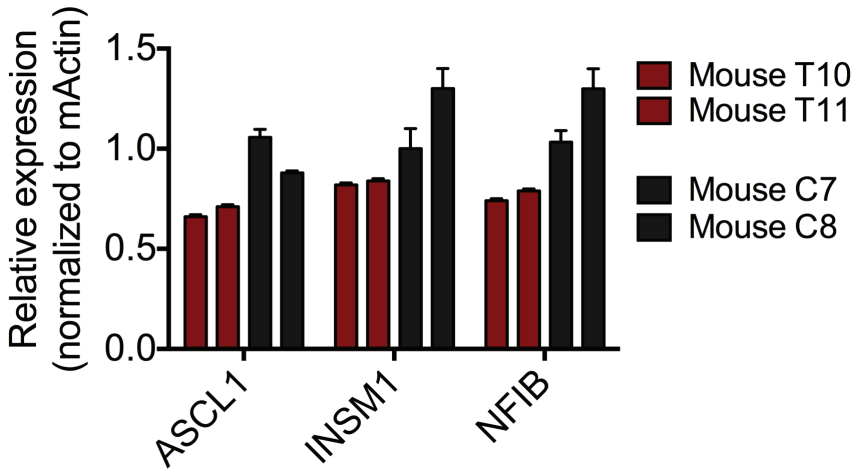

Relative expression
(normalized to mActin)
1.5
1.0
0.5
0.0
ASCL1
INSM1
NFIB
Mouse T10
Mouse T11
Mouse C7
Mouse C8

Figure S6. Related to Figure 6. Downregulation of lineage-specific transcription factors in tumors from RP mice treated with THZ1 versus vehicle

Quantitative PCR to detect expression of SCLC transcription factor genes *Ascl1*, *Insm1* and *Nfib* in RNA extracted from tumor tissue isolated from mice treated with either THZ1 (T10, T11) or vehicle (C7, C8) for 72 hours. Target gene expression is normalized to *Actin* expression and 'C7' serve as reference sample set to 'Relative expression' index 1). Data are presented as mean +/- SEM.

Table S4. Related to Figure 6. Gene enrichment analysis of THZ1-sensitive transcripts and super-enhancer associated genes

Provided as an excel file

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

High-throughput small molecule screen

Using a semi-automated platform we tested >1000 small molecule annotated library.
in 3 mSCLC cell lines (termed mSCLC1, 2, and 3), each established from a different tumor-bearing RP mouse. Murine SCLC cells were seeded in a 384-well format and treated with a concentration of 600 nM of individual compounds before evaluating cell viability after 120 hours (2 doubling times) using CellTiter-Glo Luminescent assay (Promega). The optimal cell number for the cell lines was determined prior to experiment to ensure that cells were seeded to allow for exponential growth. A compound was included for further analysis if it allowed for more than 50% reduction in cell viability as compared to control cells (no drug or DMSO) in all 3 cell lines. Further evaluation of selected compounds was performed in a 5-point 2-fold dilution series of each compound before evaluating cell viability after 120 hours using CellTiter-Glo (experimental repeat 1) or MTS-based CKK-8 assay (Dojindo) (experimental repeat 2). Mean inhibitory concentration allowing for 50 % reduction in cell viability (IC50) was calculated using GraphPad Prism 6.0 non-linear regression analysis. Compounds were ranked for potency using mean IC50 values for the 3 cell lines.

Genetically engineered mice (GEM) dosing studies and Magnetic Resonance Imaging

Upon appearance of clinical signs of disease development in mice (dyspnea, weight loss and hunched posture) Magnetic Resonance Imaging (MRI) was performed to established pre-treatment lung cancer burden. $Rb^{L/L}$;$p53^{L/L}$ (RP) mice were imaged on a 7 Tesla (BioSpec; Bruker BioSpin) with MRI protocols optimized for image requisition of pulmonary parenchyma and vessels in healthy mice. Respiratory and cardiac gating was applied to minimize motion artifacts during imaging. Animals were during the course of imaging anesthetized with 2% isoflurane IsoFlo; Abbott) in 100% oxygen via a nose cone. A total number of 24 slices of 1 mm thickness was collected providing a sufficient number to cover the lung volume. Tumor volume

per animal was quantified using 3D Slicer by manual quantification of at least 8 consecutive axial image sequences. Upon detection of tumor growth mice were randomized into treatment groups and treated with indicated drugs. THZ1: 10 mg/kg THZ1 BID (twice daily), synthesized as described in (Kwiatkowski, 2014). Vehicle: 10 % DMSO in 5 % glucose water BID. Cisplatin-Etoposide (Cis-Eto) is dosed as follow: Cisplatin 5 mg/kg 1x per week, Etoposide 10 mg/kg 3x per week. Cis-Eto is given 1-2 weeks *on* (pending weight/toxicity) hereafter 2-3 weeks *off* followed by 1-2 weeks re-treatment (pending weight/toxicity and tumor burden according to MRI). MRI was performed every second week to follow tumor volume after treatment start and weights were monitored bi-weekly during the course of treatment.

Histology and immunohistochemistry

Lungs were perfused with 10% formalin, stored in fixative overnight, and embedded in paraffin. For further staining with hematoxylin and eosin (H%E) and antibodies, sections of 5 µm were cut. TUNEL assay was performed according to instructions (Milipore). Positive cells were identified in high-power field (20x and 40x magnification) observing 3 fields per tissue section.

Human cell lines

The following human cell lines were used: GLC2, GLC3, GLC16, GLC19, GLC26, DMS53, DMS79, DMS92, DMS153, NCI-H69, NCI-H209, NCI-H82 (Berendsen et al., 1988; Carney et al., 1985; de Leij et al., 1985). All of the GLC, NCI and DMS79 lines were cultured in RPMI 1640 and all DMS (except DMS79) were cultured in Waymouth medium, all supplemented with 10% FBS (Sigma) and cultured in a humidified chamber in the presence of 5 % CO_2.

Cellular viability, proliferation and apoptosis assays

Cells were seeded in 96-well plates (0.015-0.025 x 10^6 cells/well) and exposed to media and small molecule inhibitors (THZ1 etc) at indicated concentrations and time points. Cell viability

was measured using the MTS-based CCK-8 assay (Dojindo). Absorption at 450 nm was measured 3 hours after addition of CCK-8 reagent to cells. Cell proliferation was measured using the BrdU Proliferation ELISA assay (Roche) according to instructions. Apoptosis was measured using the caspase-GLO 3/7 assay (Promega) according to instructions.

Western Blotting

Whole cell lysates were prepared by sonication in ice-cold Tris-HCL, Triton X-100 2 % supplemented with protease and phosphatase inhibitors (Promega). Tumor samples were homogenized before lysis as described above. Protein concentrations were determined using the BCA protein assay (Pierce) according to instructions. Western Blotting was performed using SDS-PAGE followed by transfer to nitrocellulose membrane using the NuPAGE PreCast Gel system (Invitrogen). Membranes were incubated with primary antibodies overnight: RNAPII (Santa Cruz Biotechnologies), RNAPII-CTD-SER2 (Bethyl Laboratories), -SER5 (Bethyl Laboratories), -SER7 (Milipore), CDK7 (Bethyl Laboratories), ACTIN (Cell Signaling), CASPASE 3 (Cell Signaling), PARP (Cell Signaling), MYCN (Santa Cruz Biotechnologies), C-MYC (Cell Signaling), INSM1 (Santa Cruz Biotechnologies), ASCL1 (BD Biosciences), NEUROD1 (Abcam), OTX2 (Abcam). Secondary antibodies were purchased from GE Healthcare.

Bio-THZ1 pulldown

Bio-THZ1 pulldown experiments followed by western Blotting of enriched proteins was performed as described in (Kwiatkowski, 2014). Briefly, cells were treated with THZ1 or DMSO for 4 hrs. Following treatment cells were washed 2-fold with cold PBS and then lysed in the following lysis buffer: 50 mM TrisHCl pH 8.0, 150 mM NaCl, 1% NP-40, 5 mM EDTA, 1mM DTT, and protease/phosphatase cocktails. Following clearance, lysates were treated with bio-THZ1 for pulldown overnight at 4 ^{o}C. Lysates were further incubated at room temperature for 3 hrs to increase the efficiency of covalent bond formation. Lysates were then incubated with

streptavidin agarose for pulldown for an additional 2-3 hrs at 4^{o}C. Agarose beads were washed 5 times with lysis buffer and then boiled in 2X SDS at 95 ^{o}C. SDS-page resolved precipitated proteins were probed for the indicated proteins.

3D matrix in vitro assay

The microfluidic tissue culture devices used in this study are described in detail elsewhere (Aref et al., 2013; Zhu et al., 2014). In short, hSCLC spheres were seeded in 3D collagen matrix and exposed to DMSO, THZ1-R or THZ1 (100 nM) for 12 hours. Phase-contrast microscopy was used to evaluate drug effect on sphere morphology.

Xenograft dosing studies

hSCLC xenograft models were established by subcutaneous (s.c.) injecting of 5 x 10^6 NCI-H69 or GLC16 cells in matrigel (Corning) into the flank of nude mice (Charles River). Approximately, 2 weeks after s.c injections, visible tumor nodules appeared. Tumors were measured with caliper and when tumor volume reached 100-200 mm^3 mice were randomized into treatment with either 10 mg/kg THZ1 BID or vehicle BID. Tumor volume and mice weights were measured weekly from treatment start.

RNA extraction and Synthetic RNA Spike-In

5 x 10^6 human SCLC cells per experiment were seeded the day before treatment. Treatments were either DMSO (control), or 25 nM or 100 nM THZ1 for 6 hours. At cell collection, cell numbers were determined by manually counting of Trypan Blue-stained cells in a hemacytometer prior to lysis and RNA extraction. Total RNA was extracted using 1ml if TRIzol reagent (Life technologies) followed by 0.2 ml Chloroform to allow for phase-separation by centrifugation. RNA was precipitated using isopropanol and further washed with 80 % Ethanol and re-suspended in nuclease-free water. Hereafter, DNAse treatment was performed on RNA

stock (Qiagen) with subsequent clean-up using RNeasy mini kit (Qiagen). ERCC Spike-In RNA Mix (Ambion) was added to total RNA in proportion to cell number to normalize for cell number as previously described (Loven et al., 2012). RNA quality was assessed using an Agilent 2100 Bioanalyzer. RNA with the RNA Integrity Number (RIN) above 9.8 was hybridized to GeneChip PrimeView Human Gene Expression Arrays (Affymetrix).

cDNA synthesis and Quantitative PCR

Total RNA was prepared as described above (*RNA extraction and Synthetic RNA Spike-In*). cDNA was synthesized using the High Capacity RNA-to-cDNA kit (Applied Biosystems) according to instructions. Quantitative PCR was performed using Tagman probes (Applied Biosystems) for indicated target genes and using GUSB (human) or ACTIN (mouse) gene as internal sample probe to correlate for inter-assay variability. Samples were prepared mixing 50 ng of cDNA, 2x RealTime Master Mix (Applied Biosystems), target probe and internal assay control probe. The reactions were run using the Step-OnePlus Realtime machine and software (Applied Biosystems). Relative quantification of expression levels was performed according to the comparative threshold cycle (Ct) method assuming equal efficiency of target and housekeeping gene probes.

Microarray Sample Preparation and Analysis

For microarray analysis, 100 ng of total RNA containing ERCC RNA Spike-In Mix (see above) was used to prepare biotinylated aRNA (cRNA) according to the manufacturer's protocol (30 IVT Express Kit, Affymetrix 901228). Briefly, total RNA undergoes T7 oligo(dT)-primed reverse transcription to synthesize first-strand cDNA containing a T7 promoter sequence. This cDNA is then converted into a double-stranded DNA template for transcription using DNA Polymerase and RNase H to simultaneously degrade the RNA and synthesize second strand cDNA. In vitro transcription synthesizes aRNA and incorporates a biotin-conjugated nucleotide. The aRNA is

then purified to remove unincorporated NTPs, salts, enzymes, and inorganic phosphate. Fragmentation of the biotin-labeled aRNA prepares the sample for hybridization onto GeneChip 3' expression arrays. Samples were prepared for hybridization using 10 µg of biotinylated aRNA in a 1X hybridization cocktail according the Affymetrix hybridization manual. Additional hybridization cocktail components were provided in the Affymetrix GeneChip Hybridization, Wash and Stain Kit. GeneChip arrays (Human PrimeView, Affymetrix 901837) were hybridized in a GeneChip Hybridization Oven at 45 degrees C for 16 hrs at 60 RPM. Washing was done using a GeneChip Fluidics Station 450 according to the manufacturer's instructions, using the buffers provided in the Affymetrix GeneChip Hybridization, Wash and Stain Kit. Images were extracted with Affymetrix GeneChip Command Console (AGCC), and analyzed using GeneChip Expression Console. A Primeview CDF that included probe information for the ERCC controls (GPL16043), provided by Affymetrix, was used to generate .CEL files. We processed the CEL files using standard tools available within the affy package in R as described in (Loven et al., 2012). The CEL files were processed with the expresso command to convert the raw probe intensities to probeset expression values. The parameters of the expresso command were set to generate Affymetrix MAS5-normalized probeset values. We used a *loess* regression to re-normalize these MAS5 normalized probeset values, using only the spike-in probesets to fit the *loess*. The affy package provides a function, *loess* normalize, which will perform *loess* regression on a matrix of values (defined using the parameter mat) and allows for the user to specify which subset of data to use when fitting the *loess* (defined using the parameter subset, see the affy package documentation for further details). For this application the parameters mat and subset were set as the MAS5-normalized values and the row-indices of the ERCC control probesets, respectively. The default settings for all other parameters were used. The result of this was a matrix of per-probeset expression values normalized to the control ERCC probes. Probe set-level expression was collapsed to log2 RefSeq transcript-level transcription by taking the probe-set with the maximum average signal across all experiments. Where possible, log2

values of biological replicates were averaged. Fold-changes were taken by subtracting average log2 DMSO signal from average log2 treatment signal. Expressed genes were those with log2(expression) > log2(100) in the corresponding DMSO sample. Expression heatmaps in Figure 4A were made using heatmap.2.

Chromatin immunoprecipitation, sequencing and analysis

Cells were cross-linked for 10 min at room temperature with 1% formaldehyde in PBS followed by 5 minutes quenching with 2.5 M glycine. The cells were then washed twice in ice-cold PBS, and the cell pellets were 'flash frozen' and stored at −80 °C.

100 µl of Dynal magnetic beads per sample (Invitrogen) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were loaded with 10 µg of H3K27Ac (Abcam cat# ab4729) antibody over night at 4^{o}C. For each ChIP analysis 2X10^{7} cells were used. Cross-linked cells were lysed in 120µl of lysis buffer (0.5% SDS, 50 mM Tris, pH 8, 10 mM EDTA, 1 Complete protease inhibitor (Roche)) per 5X10^{6} cells and chromatin was sheared in microTUBEs (Covaris) with an E210 Ultrasonicator (Covaris). Each microTUBE (120 µl of lysate) was sonicated with six treatments of 60 s each with the following settings: intensity, 5; 'duty cycle', 10%; 200 cycles per burst. That sonication sheared the chromatin into fragments 150–600 base pairs in length, with most DNA fragments 200–250 base pairs in length. The sonicated lysates were collected and then were centrifuged for 10 min at 4 °C. Supernatants were collected and four parts of dilution buffer (1.25% Triton X-100, 12.5 mM Tris, pH 8, 187.5 mM NaCl, 1 Complete protease inhibitor) were added. Sonicated lysates were cleared and incubated overnight at 4^{o}C with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Precipitated immunocomplexes were then washed, for 5 min each, as follows: once with low-salt buffer (0.1% SDS, 1% Triton X-100, 20 mM Tris, pH 8, 2 mM EDTA, 150 mM NaCl and 1 Complete protease inhibitor), once with high-salt buffer (0.1% SDS, 1% Triton X-100, 20 mM Tris, pH 8, 2 mM EDTA, 500 mM NaCl and 1 Complete protease inhibitor), twice with LiCl buffer (0.7%

sodium deoxycholate, 1% NP-40, 20 mM Tris, pH 8, 1 mM EDTA, 500 mM LiCl and 1 Complete protease inhibitor) and once with Tris-EDTA buffer (with protease inhibitor). DNA was eluted in elution buffer (50 mM TrisHCl pH 8.0,10 mM EDTA, 1% SDS). Cross-links were reversed overnight. RNA and protein were digested using RNase A and Proteinase K, respectively and DNA was purified with phenol chloroform extraction and ethanol precipitation.

Libraries were generated for ChIP samples following the Illumina TruSeq™ DNA Sample Preparation v2 kit protocol. Libraries were quantified by qPCR using the KAPA Biosystems Illumina Library Quantification kit according to kit protocols. Libraries with distinct TruSeq indexes were multiplexed by mixing at equimolar ratios and running together in a lane on the Illumina HiSeq 2000 for 40 bases in single read mode.

Illumina sequencing libraries were generated and data were processed as previously described (Lin et al., 2012) with the following modifications described here. All ChIP-seq datasets were aligned to hg19. Wiggle files for gene tracks were created using MACS (http://genomebiology.com/2008/9/9/r137) with parameters –nomodel—shiftsize=200 to extend reads artificially to be 200 bp and normalized to the millions of mapped reads per bin (RPM/bin). Inspection of input DNA suggested that the following loci were focally amplified: MYCN in NCI-H69, MYC in GLC16 and NCI-H82, and OTX2 in NCI-H82. To account for the increased amount of DNA corresponding to that region of the reference genome, background signal was subtracted from each bin in these loci to construct all MYCN, MYC, and OTX2 gene tracks (top row Figure 5D). Specifically, RPM-normalized WIGs were constructed for both H3K27ac and Input DNA from the same cell type. The RPM-normalized value from Input was subtracted from the RPM-normalized H3K27ac value for each bin for each cell type. Peaks were identified twice using MACS, once with --keep-dup=1 and once with --keep-dup=all; both sets used –p 1e-9. The union of these peaks was used to identify enhancers and capture enhancers in focally amplified regions.

Super-Enhancers were identified using ROSE (https://bitbucket.org/young_computation/rose) as described in Whyte et al, 2013, Hnisz et al., 2013, and Loven et al., 2013. Briefly, the union of peaks of H3K27ac was stitched into enhancer domains if they were within 12.5kb. Peaks were excluded if they were entirely contained within +/- 2kb from a RefSeq TSS. Enhancer domains were ranked by H3K27ac signal and a point of separation was defined between super-enhancers and typical enhancers by determining the point along the X axis at which the line Y=X was tangent to the curve. Enhancers were assigned to the RefSeq transcript whose TSS was nearest the center of the enhancer.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was performed as described previously (Subramanian et al., 2005). The dataset was converted from probe sets to gene symbols, and the gene expression signatures were analyzed using the java GSEA package. The most differentially expressed genes ranked by signal to noise ratio for each comparison were used to generate a signature for GSEA analysis. The input TFBS gene sets were extracted from the Molecular Signature Database, version 4 (MSigDBv2).

Gene Expression Profiling: Hierarchical clustering analysis was performed using pairwise complete linkage method with the Pearson correlation distance measure employing the Genepattern genomic analysis platform."Single Sample" Gene Set Enrichment Analysis (GSEA) was used to determine the degree of absolute enrichment of GO biological processes gene sets (obtained from http://www.broadinstitute.org/gsea/msigdb) in each sample within the gene expression data set. Signature values for each sample were normalized using the entire sample set. Gene expression values for a given sample were rank-normalized and rank-ordered, and an enrichment score was produced using the Empirical Cumulative Distribution Functions (ECDF) of genes in the GO signatures and remaining genes. A statistic was calculated by an integration

of the difference between the ECDFs, which is similar to the method used in GSEA but based on absolute expression instead of differential expression.

Standard GSEA (http://www.broadinstitute.org/gsea) was used to determine the enrichment of transcription factor binding sites between classes. Gene sets were obtained from MSigDB (http://www.broadinstitute.org/gsea/msigdb), a collection of gene sets that includes a database of genes sharing a cis-regulatory motif that is conserved across the human, mouse, rat, and dog genomes, representing known or predicted regulatory elements in promoters and 3'-UTRs. GSEA estimates whether the members of a given TFBS gene set are found at the top or bottom of the list, indicating they are associated with a specific phenotype (i.e., SCLC or THZ1 treatment), rather than being distributed uniformly or randomly across the list. An enrichment score (ES) is calculated to quantify the degree to which a gene set is over-represented at the top or bottom of the entire ranked list. After calculation of the scores for a collection of gene sets, an empirical phenotype-based permutation test procedure is used to estimate P-values. GSEA normalizes the ES for each gene set to account for the variation in set sizes, yielding a normalized enrichment score (NES) and a false discovery rate (FDR). The FDR gives an estimate of the probability that a set with a given NES represents a false positive finding; it is computed by comparing the tails of the observed and permutation-computed null distributions for the NES.

Functional GO term analysis

Gene ontology analysis was performed using the DAVID software version 6.7 (http://david.abcc.ncifcrf.gov).

SUPPLEMENTAL REFERENCES

Aref, A. R., Huang, R. Y., Yu, W., Chua, K. N., Sun, W., Tu, T. Y., Bai, J., Sim, W. J., Zervantonakis, I. K., Thiery, J. P., and Kamm, R. D. (2013). Screening therapeutic EMT blocking agents in a three-dimensional microenvironment. Integrative biology : quantitative biosciences from nano to macro *5*, 381-389.

Berendsen, H. H., de Leij, L., de Vries, E. G., Mesander, G., Mulder, N. H., de Jong, B., Buys, C. H., Postmus, P. E., Poppema, S., Sluiter, H. J., and et al. (1988). Characterization of three small cell lung cancer cell lines established from one patient during longitudinal follow-up. Cancer research *48*, 6891-6899.

Carney, D. N., Gazdar, A. F., Bepler, G., Guccion, J. G., Marangos, P. J., Moody, T. W., Zweig, M. H., and Minna, J. D. (1985). Establishment and identification of small cell lung cancer cell lines having classic and variant features. Cancer research *45*, 2913-2923.

de Leij, L., Postmus, P. E., Buys, C. H., Elema, J. D., Ramaekers, F., Poppema, S., Brouwer, M., van der Veen, A. Y., Mesander, G., and The, T. H. (1985). Characterization of three new variant type cell lines derived from small cell carcinoma of the lung. Cancer research *45*, 6024-6033.

Loven, J., Orlando, D. A., Sigova, A. A., Lin, C. Y., Rahl, P. B., Burge, C. B., Levens, D. L., Lee, T. I., and Young, R. A. (2012). Revisiting global gene expression analysis. Cell *151*, 476-482.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America *102*, 15545-15550.

Zhu, Z., Aref, A. R., Cohoon, T. J., Barbie, T. U., Imamura, Y., Yang, S., Moody, S. E., Shen, R. R., Schinzel, A. C., Thai, T. C.*, et al.* (2014). Inhibition of KRAS-driven tumorigenicity by interruption of an autocrine cytokine circuit. Cancer discovery *4*, 452-465.